CollegeBoard
connect to college success™

Research Report
No. 1998-7

Effects of Extended Time on the SAT® I: Reasoning Test Score Growth for Students With Learning Disabilities

Wayne J. Camara, Tina Copeland, and Brian Rothschild

www.collegeboard.com

College Board Report No. 98-7

Effects of Extended Time on the SAT® I: Reasoning Test Score Growth for Students With Learning Disabilities

WAYNE J. CAMARA, TINA COPELAND,
and BRIAN ROTHSCHILD

College Entrance Examination Board, New York, 1998

Wayne J. Camara is executive director for research and development at the College Board.
Tina Copeland is a graduate student at City University of New York.
Brian Rothschild is a graduate student at Fordham University.

Founded in 1900, the College Board is a not-for-profit educational association that supports academic preparation and transition to higher education for students around the world through the ongoing collaboration of its member schools, colleges, universities, educational systems and organizations.

In all of its activities, the Board promotes equity through universal access to high standards of teaching and learning and sufficient financial resources so that every student has the opportunity to succeed in college and work.

The College Board champions—by means of superior research; curricular development; assessment; guidance, placement, and admission information; professional development; forums; policy analysis; and public outreach—educational excellence for all students.

Additional copies of this report (item #050481642) may be obtained from College Board Publications, Box 886, New York, New York 10101-0886, (800) 323-7155. The price is $15. Please include $4 for postage and handling.

Printed in the United States of America.

Acknowledgments

The authors thank June Zumoff, who provided continued assistance throughout this research in providing the necessary data and aided in defining the appropriate groups for this study, and Cathy Wendler and Maureen Welsh, who aided in the design of this study and the review of the final report. In addition, the authors wish to thank Howard Everson, Ellen Mandinach, Brent Bridgeman, and Hunter Breland for their comments on an earlier version of this paper. This research was funded through the Joint Staff Research and Development Committee of ETS and the College Board. Robert Majoros assisted in the manuscript preparation.

Contents

Abstract

Tests administered with accommodations to persons with disabilities have been considered nonequivalent to tests administered under standardized conditions to nondisabled test takers (Sherman and Robinson, 1982). This study examined the score change patterns for learning disabled students completing extended-time administrations of the SAT® I in comparison to nondisabled students retesting under standard-time administrations. Results illustrate that learning disabled students generally perform about .5 of a standard deviation below nondisabled test takers. However, the mean score gain for learning disabled students first completing a standard-time SAT I and retesting under an extended-time SAT I is more than three times as large as the mean score gain for both nondisabled students testing under standardized conditions and learning disabled students testing with extended time on both occasions.

Introduction

Each year thousands of students with disabilities receive special accommodations when completing standardized educational tests. Accommodations include providing people to read aloud instructions and test questions; cassette, large print, and Braille forms of tests; individualized administration; and extended time in completing tests. Such accommodations are designed to compensate for the disability by removing an extraneous source of difficulty the test may impose on some students. That is, failure to adjust the test or testing conditions in some way would be unfair to a student with a disability.

Standardization is essential in testing because it attempts to ensure that all students have an equal opportunity to demonstrate their skills and knowledge through testing conditions and content that are consistent for all test takers (Willingham et al., 1988). Accommodations are not designed to remove all standardization, but rather to make appropriate adjustments, given a student's disabling conditions, that will provide him or her with an equal opportunity to demonstrate his or her skills and knowledge. As such, accommodations should increase comparability of information from the test for test takers by reducing or eliminating irrelevant sources of difficulty (e.g., visual, motor). Testing accommodations should level the playing field for a person with a disability. For example, reading a test to a person who is blind would be a reasonable accommodation. The test performance of the person who is blind would be judged by the same criteria as his or her sighted peers, but the reader would accommodate the person's blindness (Eyde et al., 1994).

However, research has generally found that results from testing with accommodations are not comparable to results from standard administrations (Sherman and Robinson, 1982). This is particularly evident for students with learning disabilities. In separate studies of the SAT, GRE, and LSAT, researchers found admission tests for learning disabled students substantially overpredicted their resulting performance in school (Braun, Ragosta, and Kaplan, 1988; Wightman, 1993). Willingham et al. (1997, p. 185) states, "the primary source of noncomparability that is associated with test scores is the extended time…several types of evidence suggest that in some cases the extra time is excessive and results in noncomparable scores." Research conducted on the ACT also suggests that the estimates of college success are not the same for individuals who request to take the ACT assessment under nonstandard conditions and those who take the test under standardized conditions. There is a strong tendency for freshman grade-point averages (GPAs) to be overpredicted for disabled students (ACT, personal communications, January 25, 1995). Of the predictive validity study conducted on the LSAT, Wightman (1993, p. 53) notes, "results [for learning disabled students] demonstrate a strong tendency toward overprediction…currently, the scores cannot be relied upon to provide indications of first-year performance in law school to the same extent that scores earned by students at regular LSAT administrations can be." One of the most challenging issues facing test administrators is how to interpret scores for students with disabilities who test with accommodations.

These issues are particularly salient for admission tests where colleges and universities compare performance among students to aid in admission decisions. Similarly, students, their parents, teachers, and counselors use results from admission tests to make norm-referenced comparisons of students that may assist in selecting colleges to which students will apply.

The number of students reporting disabilities who receive accommodations on the SAT I has doubled in five years, with annual increases averaging 14 percent (Welsh, 1995). The number of requests for accommodations on the ACT has increased by two-thirds in a four-year span (Dana and Ziomek, 1996). Among students with disabilities, those with learning disabilities represented nearly 90 percent of all students receiving accommodations on the SAT I in 1995-96, with the largest annual increases in number and proportion of requests (ETS, 1996). Learning disabilities is a broad classification that encompasses several more specific forms of mental disability. Classification of learning disability may include dyslexia, attention deficit/hyperactivity disorder, and other specific mental disorders. The great variety of learning disabilities and the fact that they often occur in variable combinations, resulting in multiple sensory and processing problems, may often require multiple accommodations.

Extended-time or untimed testing is a frequent accommodation for students with learning disabilities. Additional time can provide a person with a visual sequencing disturbance more time to properly see letters, numbers, or objects in the correct order, and to read and perform mathematical computations (Eyde et al., 1994). Additional time may also permit persons with dyslexia to read a test more slowly to improve comprehension.

The Rehabilitation Act of 1973 and the ensuing federal regulations promulgated under Section 504 of the act extended civil rights protections to persons with disabling conditions. The Americans with Disabilities Act (ADA) of 1990 defined "disability" broadly and requires test publishers and administrators to provide reasonable accommodations, including adjustments or modifications to exams and administration, based on an individual's disabling condition (Asquith and Feld, 1992). This most recent legislation requires individual accommodations, on a case-by-case basis, which would best meet the needs of the individual student. A 20 percent increase in time may be viewed as an adequate accommodation for one learning disabled student, while extended time of 50 percent or 200 percent may be requested for other students with the same disability and level of severity.

Various options for extended time are offered as accommodations by all major admission testing programs. Typically, unlimited timing is not approved (Mehrens, 1997), but timing policies do provide individual test takers with adequate time based on the nature and type of their individual disability. The ADA's emphasis on individualized, case-by-case accommodations appears to prohibit many of the psychometric solutions proposed by researchers and policymakers for increasing score comparability (AERA, APA, and NCME, 1985; Mehrens, In Press; Sherman and Robinson, 1982; Willingham et al., 1988). The individual accommodation provision greatly reduces the opportunity to conduct research examining the comparability of scores as more common time limits for similar disabilities are reduced. In addition, the incentive to conduct research that would examine comparability of scores for common standardized accommodations (e.g., a given increase in testing time for specific types of disabilities) is greatly reduced because such procedures may not be defensible under the ADA. The ADA requires accommodations be granted on a case-by-case basis (Mehrens, 1997).

Ragosta and Wendler (1992) investigated the effects of timing in admission testing and concluded that the goal is to have students with disabilities finishing at rates reasonably comparable to nondisabled test takers, somewhere between time and a half and double time. The assumption is that it is reasonable to have a similar percent of disabled and nondisabled test takers completing a test. The current policy of admission tests to permit varying amounts of extended time, based on the individual needs and requests of the test taker, has resulted in differences in student completion rates and the amount of time to complete the test (Willingham, 1987). The effects of extended time have become increasingly important as this is the only accommodation for more than 7 of 10 students completing the SAT I.

Centra (1986) specifically examined the effects of extended time on score changes for students with disabilities completing the SAT between 1979-82. He studied students who took both an extended-time and standard-time SAT. For this group of test takers, the scores from the standard-time administration were well below the national mean (by more than .5 standard deviation), while the scores from the extended-time administration were only somewhat lower than average (less than .3 standard deviation). Two methods were applied to examine the score differences of disabled and nondisabled test takers. First, Centra examined mean score differences for disabled students who first took an extended-time SAT and disabled students who first took a standard-time SAT. Averaging the differences between these two groups of test takers is one method of attempting to control for growth, practice, and error effects associated with score change over time. A second method was to subtract the score gain for disabled students taking a standard-time test followed by an extended-time test from the mean junior-to-senior score change data from the national administration. He reported score changes of 37 and 30 points, respectively, for each method on the SAT Verbal scale, and score changes of 38 and 33 on the SAT Math scale. Centra (1986) also noted that score gain increased in a direct relationship to the amount of extended time taken by the learning disabled students.

However, many changes have occurred in the nearly two decades since students in Centra's study completed the SAT, including passage of the ADA and the reauthorization of the Individuals with Disabilities Education Act (IDEA) in 1997. While the ADA greatly expanded protections for persons with disabilities in pre-employment, employment, and assessment, the IDEA calls for a broader assessment of children's performance using multiple measures, including the authority to modify and create assessment tools and procedures that can accurately determine any special needs of children with disabilities.

Ziomek (1997) investigated the score change on the ACT composite scale. He noted that "students who tested at least twice under special conditions had a higher average ACT Composite gain…as compared to non–special needs students who tested at least twice under regular conditions, [and] special needs students who tested under regular conditions followed by special testing conditions benefit substantially" with a gain of 3.2 points (pp. 2–3). Eighty-six percent of disabled test takers who first completed a standard-time ACT had score increases, as compared to 27 percent of disabled students who tested first with accommodations, and 58 percent of students who tested twice with special accommodations. Results are presented separately for students with dyslexia, attention deficit/hyperactivity disorder, and learning disabilities.

Research on the effects of extended time for nondisabled test takers has generally found that additional time produces small to moderate score gains on both essay and multiple-choice tests (Evans, 1980; Evans and Riley, 1972, 1973; Klein, 1981; Wild and Durso, 1979). However, there is no simple relationship between student performance and test speededness. Among different groups of test takers, the relationship varies with the type of question, by content area, or skill; at the individual level, there are style and strategy differences that affect student performance and speed. Score differences become more pronounced when tests are administered under highly speeded conditions. Generally, test timing has not differentially impacted the performance of any ethnic minority or racial group in these studies (Evans, 1980), and findings have been inconsistent concerning gender differences associated with test speededness (Bridgeman and Schmitt, 1997).

The purpose of this study was to replicate the earlier study conducted by Centra (1986) on the revised SAT I with learning disabled students, who have represented an increasing proportion of test takers in recent years, and to examine the residual score gain associated with extended-time forms at each score level for the initial test. The earlier study did not examine score change by initial score, and the reported results may be deceptive for students with initial scores above or below the 300–400 range examined. In addition, this study also examined the score gain for disabled students completing two extended-time administrations, as well as nondisabled students who completed two standard-time forms of the SAT I. Finally, a larger and more representative sample of learning disabled students is employed in this study than was available to Centra (1986).

The SAT I is an enhanced revision of the former SAT and was first introduced at the March 1, 1994 administration. The SAT I includes some new question formats, more emphasis on critical reading skills, and fewer items across more sections, but total testing time is still three hours. Testing time for the math and verbal sections each increased from 60 minutes to 75 minutes, with seven fewer items in the verbal section and no change in the number of math items. The following year a recentered scale was introduced, and all scores reported in this study, including tests administered prior to April 1995, are reported on the recentered scale.

The SAT I comprises seven sections. Each standard form of the SAT I has two 30-minute math sections and verbal sections, one 15-minute math section and verbal section, and one 30-minute pretesting or equating section. Swineford's (1956) guidelines are used for assessing the speededness in SAT I test analysis: (1) 80 percent of examinees should have an opportunity to reach the last question on a test, and (2) all examinees should be able to reach 75 percent of items. The SAT I and many similar admission and cognitive ability tests are not meant to be speeded tests for any group (Evans, 1980). Time limits are imposed more for efficient administration (to ensure some students are not waiting for long periods of time while other students are still working on the test) and test security than to evoke a speed factor.

Data from seven recent forms of the SAT I indicate that the average number of questions not reached by students was between .3 and 1.0 for verbal sections and .2 and 1.5 for math sections. In addition, 94 percent to 100 percent of students completed 75 percent of items in verbal sections while 97 percent to 100 percent of students completed 75 percent of items in math sections (College Board, 1997). Students receiving extended time do not have such constraints—instead, sufficient time is provided to each student so that they can consider each question. However, students completing an extended-time administration may engage in different types of self-regulatory behaviors than students completing a standard-time administration of the test, particularly in pacing themselves, because time limits for each section are removed. Table 1 provides a comparison of timed conditions for students completing standard-time and extended-time forms of the SAT I.

This study was restricted to students with learning disabilities for several reasons. The number and proportion of students with learning disabilities requesting accommodations on the SAT I have been increasing annually at a rate substantially faster than for students with any other type of disability, and represents approximately 90 percent of students receiving accommodations on the test. In addition, 67 percent of students

TABLE 1

Comparison of Timing Conditions Under Standard-Time and Extended-Time Administrations

Standard Time	*Extended-Time Accommodation*
Students have three hours to complete the test	Students have no fixed limit—but on average take 4.5 hours to complete the test*
Students have separate time limits of 15–30 minutes for each of seven sections	No section time limits are imposed
80 percent of test takers reach all items	Each test taker reaches all items
Examinees complete an average of 75 percent of the items	Each test taker reaches all items

*Students testing under PLAN B specify in advance that they will complete the SAT I in time and a half (an additional 90 minutes), and the only accommodations are extended time and large print, if needed.

with learning disabilities used a regular-print test and had no accommodation other than extended time.

The majority of students with other types of disabilities receive multiple accommodations such as cassette, Braille, or large-print forms; extended time; readers; etc. It is difficult to estimate the effects of extended time for students with visual, hearing, physical, and other types of disabilities, both because so many of these students do receive multiple accommodations, which would introduce additional uncontrolled sources of variation to test conditions, and because these populations of students completing the SAT I each year are extremely small in comparison to students with learning disabilities.

Method

This study examined the scores of all students completing the first 13 administrations of the revised SAT I: Reasoning Test between March 1, 1994, the date the revised test was introduced, and December 1995. Approximately 2 million students completed the SAT I during this period of time. Of these students, over 4 percent reported having a disability when registering for the test, and 32,277 students (1.58 percent) were tested under nonstandardized conditions, including various accommodations such as extended time, Braille or cassette formats, and readers.

First, we were interested in score change for only those students who completed the SAT I during the spring of their junior year and repeated the test in the fall of their senior year. This would occur for students who completed the SAT I as juniors in March, May, or June of 1994 and again as seniors in October, November, or December of 1994, or students who completed the SAT I as juniors in January, April, May, or June of 1995 and again as seniors in October, November, or December of 1995. Students in this study were expected to graduate from high school in 1995 or 1996. When students completed more than one SAT I administration in either time period, the first administration was used. All SAT I scores were reported on the recentered scale.

Next, we identified four groups of repeat test takers from the SAT I history file; three groups of students with learning disabilities who had various patterns of repeating the SAT I with one or more extended-time administrations, and the fourth group of students with no reported disabilities who repeated the SAT I under standardized conditions. Additional students with learning disabilities who received multiple accommodations were also eliminated from the samples so results concerning extended time would not be confounded with other accommodations (e.g., reader, large-print form).

Of the two million students completing the SAT I between these dates, 51 percent of nondisabled students completed the SAT I only once and 13 percent of students did not have the designated fall to spring repeater patterns (e.g., tested as sophomores in spring 1994, never tested as seniors). Both groups of students were eliminated from the sample. A final group of 706,537 students without disabilities completed two or more standard-time administrations of the SAT I between the specified dates and was designated as Group 1 in the study. Typically, the second testing occurred about 8 months from the initial testing, but ranged from 4 months to 11 months in some extreme instances.

Similarly, 32,277 students completed the SAT I with accommodations during the first 13 administrations. In this group, 15,573 completed only one nonstandard SAT I or did not have the designated fall to spring repeater patterns and were also eliminated from the sample. Of the remaining 16,704 students, 15,444 students (92.5 percent) reported having a learning disability. A final sample of 9,112 students (59 percent) with learning disabilities who received an extended-time accommodation and had met specified conditions for repeating the SAT I were retained in the sample. A small number of students meeting all conditions were not included in the study because timing data was missing for one or more sections of their extended-time testing.

TABLE 2

Students With Disabilities Completing the SAT I Selected for Final Sample

Condition	*Students in the Sample*	*Eliminated From Sample*
Testing Between 4/1/94–12/96	32,277	—
Completed two or more administrations of SAT I during designated dates	16,704 (51.8)	-15,573 (48.2)
Students with a visual disability		-397 (2.4)
Students with a physical disability		-376 (2.3)
Students with a hearing disability		-131 (.8)
Missing data		-365 (2.1)
Students with a learning disability	15,444 (92.5)	-1,260 (7.5)
Students with 1 or more extended-time administrations of the SAT I	9,112 (59.0)	-6,323 (40.9)

Fifty percent of students with learning disabilities who repeated the SAT I more than once took a standard-time administration of the test. About 83 percent of these students completed the standard-time form prior to an extended-time form. Table 2 provides the distribution of students with disabilities selected for the study. These students were then placed into one of three groups (either Group 2, 3, or 4) based on the order of completing the SAT I with or without extended time as an accommodation.

The final four groups selected for the study are described below:

1. Group 1: 706,537 students with no reported disabilities who completed two or more standard-time administrations of the SAT I and who never took the SAT I with special accommodations;
2. Group 2: 4,479 students with a learning disability who completed two or more extended-time SAT I administrations of the SAT I;
3. Group 3: 3,836 students with a learning disability who completed a standard-time administration of the SAT I and retested with an extended-time SAT I; and
4. Group 4: 784 students with a learning disability who completed an extended-time SAT I followed by a standard-time administration of the SAT I.

A slight majority of students with learning disabilities will complete only one extended-time SAT I. However, of those learning disabled students taking the test more than once, about 42 percent will first take a standard-time administration followed by an extended-time administration, while 49 percent will complete two or more extended-time administrations. Only 784 students (8 percent) will complete a standard-time administration after first completing the SAT I with an accommodation of extended time. This latter pattern of test taking is fairly rare.

Results

The means and standard deviations for the four groups of students completing the SAT I on two or more occasions, along with the mean gain scores, are presented in Table 3.

Lyu and Lawrence (1998) report that the majority of students complete the SAT I only once, yet 48 percent of students will repeat the SAT I on two or more occa-

TABLE 3

Mean SAT I Scores and Standard Deviations for Students Repeating the SAT I

	Junior-year Math Mean	*Senior-year Math Mean*	*Math Gain*	*Junior-year Verbal Mean*	*Senior-year Verbal Mean*	*Verbal Gain*
Group 1: Nondisabled Students	516.2	528.0	11.8	510.4	523.3	12.9
(standard – standard) n = 706,537	(96.1)	(100.4)	(47.1)	(95.0)	(98.4)	(47.8)
Group 2: LD Students	459.8	472.2	12.4	464.4	479.6	15.3
(extended – extended) n = 4,479	(100.9)	(105.9)	(51.5)	(97.6)	(100.3)	(53.8)
Group 3: LD Students	452.2	490.3	38.1	449.1	493.7	44.6
(standard – extended) n = 3,836	(100.5)	(105.3)	(57.5)	(97.0)	(99.3)	(58.2)
Group 4: LD Students	442.6	436.5	-6.1	443.6	434.9	-8.6
(extended – standard) n = 784	(99.3)	(93.2)	(53.4)	(97.5)	(93.4)	(57.4)

sions. The most common repeat pattern is for students to complete the SAT I twice, once each in their junior and senior years. Thirty-eight percent of total SAT I test takers follow this pattern of repeat testing. This group of test takers has the highest SAT I verbal mean, and both their verbal and math scores exceed the national means for all SAT I test takers. Results from the current study are consistent with these findings, showing mean scores of 528 and 523 for the senior SAT I Math and Verbal scores for students completing the SAT I in their junior and senior years.

Repeating the SAT I, or other similar standardized educational tests during students' junior and senior years in high school, generally results in an increase in a student's score. This growth could result from a variety of factors (i.e., increase in student abilities, practice efforts, familiarity with the test) (Nathan and Camara, 1998).

In this study, the score gain for each group was the difference between students' junior year, spring semester, SAT I score and their senior year, fall semester, SAT I score, as illustrated in equation 1:

EQUATION 1:

Mean score gain = (Senior-year, fall SAT I – Junior-year, spring SAT I)/n

Next, the effects of extended time on score gains for students with learning disabilities were estimated. There are several methods that attempt to estimate score gains for disabled test takers by adjusting for gains for growth, practice, and error effects. Centra (1986) used two methods. First, he averaged the score gain for students completing the standard-time SAT prior to the extended-time SAT and the score gain for students completing the extended-time test prior to the standard-time test. He reasoned that averaging the score gains for students completing extended-time and standard-time tests in different order cancels out the effects of growth, practice, and error.

EQUATION 2:

Effects of extended time = Mean score gain (Group 3) – Mean score gain (Group 4)/2

Using mean score gains reported in Table 3 for Groups 3 and 4 reported in equation 2 results in residual gains and estimated effects of 44.2 (38.1–[-6.1]) and 53.2 (44.6 – [-8.6]) points on math and verbal scales respectively.

However, this method likely overestimates the effects of extended time. Learning disabled students taking extended-time tests first may differ in meaningful ways from learning disabled students taking the standard-time test first. Differences in the type and degree of disability, adjustment to the disability, need for accommodation, familiarity with the test, and ability level are likely to be found between groups that would affect estimates based on this method.

Instead, this study employed an alternative method to estimate score gains for disabled test takers, which was also employed by Centra (1986). The method used in this study is to subtract the average standard-time junior to extended-time senior SAT I score gain found for students with learning disabilities (Group 3) from the average junior-year to senior-year score gain for students without disabilities testing with standard time (Group 1):

EQUATION 3:

Effects of extended time = Mean score gain (Group 3) – Mean score gain (Group 1)

Limitations apply to this and all similar methods used to estimate the effects of extended time on the score gain of students with disabilities because we are comparing two groups of test takers who differ by the presence of a disability in one group. In this study Groups 1 and 3 differ in terms of a learning disability and possibly in other significant ways that may affect the extent of the score gain. An accommodation of extended time is designed to even the playing field for students with disabilities—making their score more comparable to standard-time scores of students without a disability. Therefore, any score gain for Group 3 introduces two additional sources of variation beyond growth, practice, and error effects common across all Groups: (1) gains resulting from the extended time and more liberal testing conditions (e.g., providing additional time to complete all items on the test and review responses, absent a speeded component present on standard-time tests, provides an advantage) and (2) gains resulting from an accommodation (i.e., extended time) that should reduce the effects of the disability on test performance. Accommodations are ideally designed to address the latter, while any gain associated with the former introduces construct-irrelevant variance and measurement error.

Extended time is provided for learning disabled students because even minimally speeded tests are assumed to differentially penalize these students. Score gains achieved with extended time may better represent a disabled student's true ability than scores achieved under standard-time conditions. Therefore, all estimates of the effects of extended time using score gain comparisons between different groups may be somewhat crude estimates of the effects of extended time.

Using equation 3 results in a 31.7 (44.6 – 12.9) score increase in the verbal scale and a 26.3 (38.1–11.8) score increase in the math scale associated with extended time. This method results in lower mean score gains in both this study and Centra's study (1986) of disabled test takers in 1979-82. Results from this method show score gains that are slightly higher than the 30 verbal gain and lower than the 33 math gain reported by Centra. Unfortunately, Centra did not report mean or median score gain, but instead reported the score gain for students whose initial standard-time SAT score was in the 300–400 range. Disabled students most typically score in this range with initial testing. Applying equation 3 to students with initial SAT I scores between 300–400 results in residual score gains of 21.3 (44.9–23.6) and 9.6 (31.1–21.5) on the verbal and math scales, respectively.

The mean differences across students at all score levels do not control for the effects of the initial score on the test. The mean score gain differs substantially depending on one's initial score. The higher one's initial SAT I score as a junior, the more likely subsequent scores will be lower or have a smaller increase. Similarly, the lower a student's initial score as a junior, the more likely subsequent scores will be higher with retesting (College Board, 1997). Therefore, the initial test score should be controlled for when estimating the effects of extended time on score gain for individual students or groups of students. Tables 4–5 report the mean senior-year SAT I score for Group 3 based on students' junior-year scores. These tables illustrate the patterns of increases found across different score ranges. For example, learning disabled students with initial scores of 200–270 on a standard-time test had mean score increases of about 100 points on the verbal and math scales, respectively. However, these results must be contrasted with score increases for nondisabled students with initial scores in the same range.

Tables 6–7 provide similar retesting data for nondisabled students. Nondisabled students with initial scores of 200–270 had similar increases with retesting; the mean score with retesting was identical for both groups in math and only 27 points lower for nondisabled students in verbal. The restricted range of scores for these students probably is responsible for the extreme gains resulting from retesting. Students with initial scores of 280–320, and all subsequent ranges, had more moderate score gains when retesting with extended time.

Tables 4–5 can be used in combination with Tables 6–7 to compare the typical score gain for students in Groups 3 and 1. The mean senior-year score for students in Group 3 can be compared to the mean senior-year score for nondisabled students in Group 1 to provide an estimate of the typical difference between groups. Again, these differences may be partially attributed to the accommodation of extended time, but other differences between these populations could also affect discrepancies between mean senior-year scores. Comparable score changes are reported for students in Groups 2 and 4, controlling for initial SAT I score in the Appendix.

TABLE 4

Percentage of Learning Disabled Students With Score Changes on the SAT I Verbal Section When Testing With Standard Time in Junior Year and Retesting With Extended Time in Senior Year (Group 3) (*n*=3,836)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				1	3	8	7	17	19	21	24	337	122
280–320		1	2	3	5	14	19	18	15	13	10	357	199
330–370	1	1	2	3	7	13	22	20	17	10	4	394	476
380–420		1	2	4	7	15	27	17	14	8	5	442	712
430–470		1	2	4	7	17	21	21	14	8	5	492	822
480–520		1	1	2	8	16	24	20	16	7	5	545	775
530–570			1	4	7	17	26	22	14	5	4	591	413
580–620			1	3	12	16	22	20	17	8	1	640	244
630–670			3	5	16	7	18	28	13	3	3	681	62
680–720			4	3	34	18	26	11		7		706	27
730–770				17	33	17	33					742	6
780–800													0

Table 5

Percentage of Learning Disabled Students With Score Changes on the SAT I Math Section When Testing With Standard Time in Junior Year and Retesting With Extended Time in Senior Year (Group 3) (*n*=3,836)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				3	4	7	12	17	17	15	25	324	94
280–320		1	2	5	10	9	22	21	17	9	4	343	204
330–370		1	2	6	11	15	20	20	14	8	3	385	446
380–420	1	1	2	4	14	20	23	18	10	5	2	425	774
430–470		1	1	5	10	18	24	20	11	6	4	486	831
480–520			1	4	9	17	35	20	14	8	6	545	673
530–570			1	3	8	15	23	20	17	7	6	597	426
580–620			2	2	8	21	20	22	14	7	4	641	250
630–670			1	5	12	17	28	18	9	7	3	685	104
680–720		3		12	12	34	15	15	3	6		711	33
730–770													0
780–800							100					800	1

Table 6

Percentage of Nondisabled Students With Score Changes on the SAT I Verbal Section From Junior to Senior Year (Standard Time) (Group 1) (*n*=706,537)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				3	4	11	14	18	18	15	17	310	6,195
280–320		1	3	4	8	14	20	21	16	8	5	342	12,378
330–370	1	1	1	6	10	18	23	21	12	5	2	380	35,411
380–420		1	2	6	13	22	25	19	8	3	1	421	73,985
430–470		1	2	6	16	24	25	17	7	1	1	465	118,193
480–520		1	2	7	17	27	25	14	5	2		510	153,324
530–570		1	2	8	19	26	25	13	4	2		558	131,870
580–620		1	2	8	19	27	22	12	5	2		608	93,761
630–670		1	2	10	19	27	22	12	5	2		656	52,847
680–720		1	4	11	21	25	20	11	6	1		712	21,295
730–770		2	7	14	23	25	19	10				738	5,870
780–800	1	3	9	19	23	36	9					787	1,408

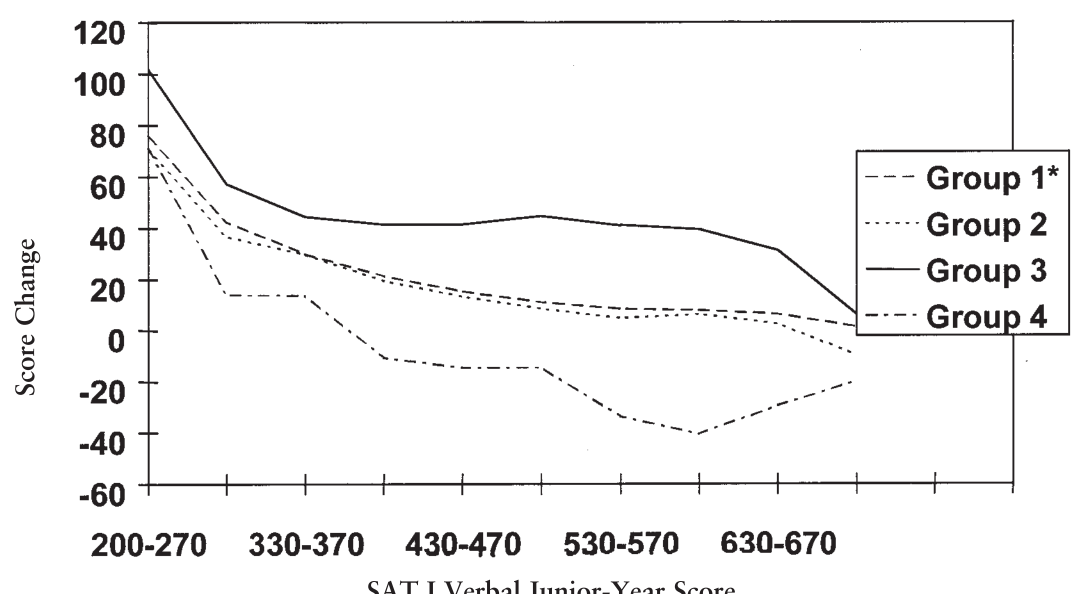
Table 7

Percentage of Nondisabled Students With Score Changes on the SAT I Math Section From Junior to Senior Year (Standard Time) (Group 1) (n=706,537)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				2	3	6	11	17	21	19	21	324	4,284
280–320		1	2	4	7	14	20	22	18	8	4	344	10,252
330–370	1	1	2	6	11	20	24	20	10	4	1	374	32,407
380–420		1	3	6	14	24	26	17	6	2	1	416	76,404
430–470		1	2	6	16	26	26	15	6	1	1	464	120,810
480–520		1	1	7	17	27	24	15	6	2		513	141,527
530–570		1	2	7	19	26	23	14	6	2		560	124,792
580–620		1	2	9	18	27	23	13	5	2		608	97,302
630–670		1	3	9	20	28	21	12	4	2		654	61,483
680–720		1	4	12	23	26	20	8	5	1		698	27,581
730–770	1	2	7	17	25	22	17	9				733	7,123
780–800	2	5	12	26	19	31	5					750	2,572

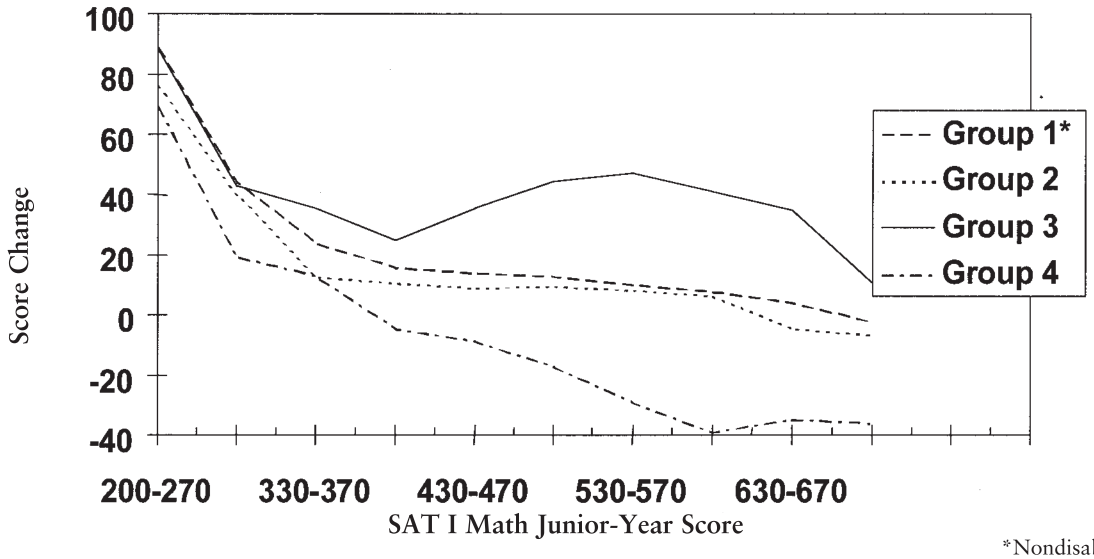
The score gains for three groups of learning disabled students (Groups 2–4) used in these analyses were compared to mean gains of nondisabled students who took the standard-time version in their junior and senior years. The means and standard deviations for all four groups of students who took the SAT I in their junior year and in their senior year are presented in Table 3. Figures 1 and 2 illustrate the mean score change with retesting for each of these four groups controlled by the initial SAT I score.

The 706,537 students completing the SAT I on two or more occasions within the designated dates had higher initial and final SAT I means than students with disabilities and means, for both the initial junior and subsequent senior administrations of the test exceeded the national means in both verbal and math. The av-



*Nondisabled test takers

Figure 1. SAT I verbal junior to senior score change for learning disabled and nondisabled test takers.



*Nondisabled test takers

Figure 2. SAT I math junior to senior score change for learning disabled and nondisabled test takers.

erage score gain from junior-spring to senior-fall administrations of the SAT I were 11.8 in math and 12.9 in verbal for these students. The mean score gain in math is slightly lower than those found for college-bound seniors in 1996 who repeated the SAT I on two occasions, while the verbal gain score is nearly identical. In that study, Lyu and Lawrence (1998) found increases of 16 and 13 points, respectively, in math and verbal scores between junior- and senior-year testing. One small difference in the samples used in these studies may partially explain the smaller mean math gain scores found in the present study. The present study included students who completed *two or more* administrations of the SAT I between designated dates, and used the first score from the spring of the student's junior year and fall of his or her senior year if multiple administrations were completed. This results in a slightly different sample than the earlier study, which was restricted to students who only completed one junior-spring and fall-senior SAT I administration.

The mean gain score was computed by subtracting each student's initial score from his or her junior-spring SAT I from his or her senior-fall test and averaging across all students in each group. The SAT I mean scores

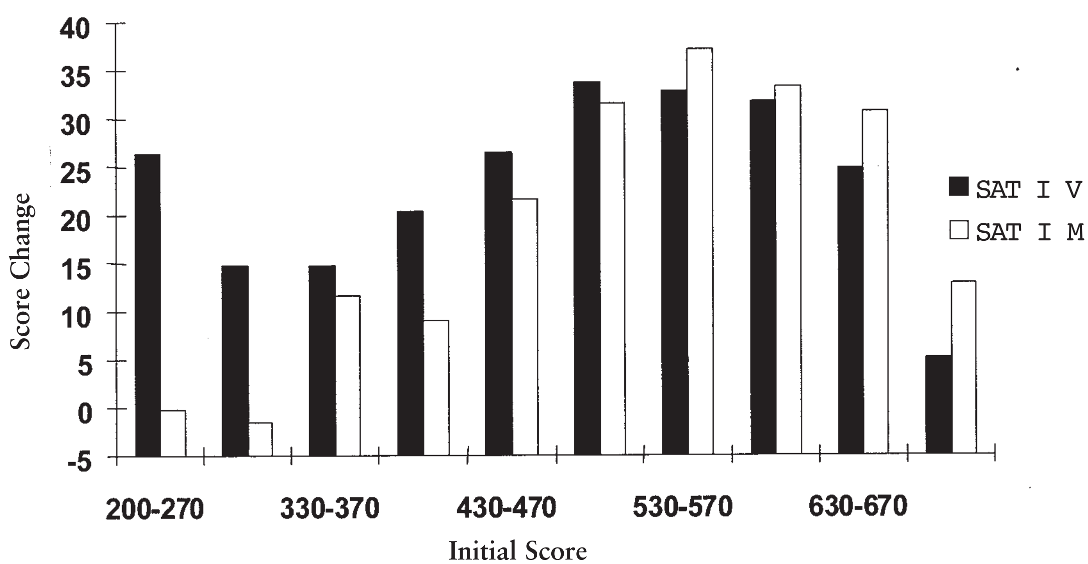


Figure 3. SAT I score change for students with disabilities. Subtracting growth for disabled students (standard time to extended time) from growth for nondisabled students under similarly timed conditions.

for students with learning disabilities were approximately .5 to .6 of a standard deviation lower than mean scores of students without disabilities. The gain scores differed substantially among groups of students with disabilities who completed standard-time or extended-time administrations of the SAT I in various orders.

Among the students with learning disabilities, the mean gain scores for students completing two extended-time administrations (Group 2) were most similar to the gains achieved by nondisabled students. Students who first completed a standard-time SAT I and then repeated with an extended-time administration (Group 3) achieved gains that were about three times greater than students in Groups 1 or 2. Figure 3 illustrates the mean score change for this group given their initial SAT I score.

The relatively small sample of learning disabled students completing an extended-time SAT I in their junior year and retesting as seniors with a standard-time administration actually had mean decreases in test scores. However, caution is required in using these findings to interpret score changes for individuals and groups. The standard deviation of score gains is quite high and the standard error of measurement across all groups ranged between 29–33 points and 30–33 points on the verbal and math scales, respectively.

Effects of Additional Time on Score Gains

The amount gained on the verbal and math sections of the SAT I increased somewhat proportionally to the amount of extra time the students used on the extended-time SAT I. Table 8 examines the effect of the amount of additional time on mean score gains for the SAT I students in Group 3 who first completed a standard-time administration followed by an extended-time test in their senior year.

The verbal and math sections of the SAT I each permit 75 minutes for standard-time administrations. Students in the first row of Table 8 requested and received an accommodation of extended time, but were able to complete each math and verbal section in the standard time permitted. These students were provided with extended time but did not use the accommodation. Their score gains of 19.8 and 23.7 points on the math and verbal scales, respectively, are larger than score gains for students without disabilities, but remain substantially lower than gains for students requiring up to time and a half (38 additional minutes; receiving up to 113 minutes for a section timed at 75 minutes) or double time (75 additional minutes; receiving up to 150 minutes for a section timed at 75 minutes) on the SAT I. Students with learning disabilities who completed a section with up to time and a half are in the second row of Table 8 and experienced mean score gains of 36.3 and 38 points, respectively, on the math and verbal

Table 8

Mean SAT I Scores and Standard Deviations for Learning Disabled Students in Group 3

Extended Time in Minutes	*Junior-year Math Mean*	*Senior-year Math Mean*	*Math Gain*	*Junior-year Verbal Mean*	*Senior-year Verbal Mean*	*Verbal Gain*
≤ = 0	430.6	450.4	19.8	438.1	461.8	23.7
	(91.8)	(104.4)	(56.1)	(91.7)	(104.7)	(61.1)
	n = 1,564	*n* = 1,564	*n* = 1,564	*n* = 1,281	*n* = 1,281	*n* = 1,281
1–38	458.5	494.8	36.3	447.0	485.0	38.0
	(92.7)	(99.8)	(50.3)	(94.8)	(98)	(52.4)
	n = 1,408	*n* = 1,408	*n* = 1,408	*n* = 1,329	*n* = 1,329	*n* = 1,329
39–75	477.0	579.9	102.9	464.6	524.8	60.2
	(86.6)	(89.0)	(53.1)	(86.6)	(95.4)	(50.9)
	n = 574	*n* = 574	*n* = 574	*n* = 827	*n* = 827	*n* = 827
≥ 76	489.9	559.0	69.1	460.8	545.6	84.8
	(77.2)	(99.9)	(56.5)	(79.1)	(88.3)	(55.8)
	n = 281	*n* = 281	*n* = 281	*n* = 390	*n* = 390	*n* = 390
Total	452.2	490.3	38.1	449.1	493.7	44.7
n = 3,827*	(92.5)	(108.7)	(57.5)	(91.7)	(102)	(58.2)
	n = 3,827	*n* = 3,827	*n* = 3,827	*n* = 3,827	*n* = 3,827	*n* = 3,827

*Data on time required to complete sections of the SAT I were missing for 9 students from Group 3.

scales. Students receiving up to double time on the math scale achieved the highest mean score gain of over 102 points when retaking the SAT I, while the verbal score gain was about 60 points. Only a small proportion of students retested with more than double time (i.e., over 150 minutes) on the SAT I, and their score gains were 69.1 and 84.8 points on the math and verbal scales, respectively. Table 8 also illustrates that, on average, students offered extended time on the SAT I required more time for completing the verbal section than the math section. Also, students with initially higher SAT I scores on the standard-time administration took more time on the subsequent extended-time administration.

The correlations between standard-time and extended-time SAT I scores were computed for students in each group to determine if student rankings are relatively stable across administrations. The correlations between junior-spring and senior-fall SAT I scores, reported in Table 9, range from .820 to .887, and are significant (p >.01) for all groups of students. Interform correlations of .887 and .878 for nondisabled students were well within the range of .86 to .89 reported by Donlon (1984) for SAT junior and senior retesting spanning a 12-year period. Single-form reliability of the SAT I is higher than interform correlations, ranging from .91 to .93 for verbal and .91 to .94 for math scales (College Board, 1997).

Interform correlations for the learning disabled students completing two extended-time tests were similarly high (.877 and .852 for verbal and math scales, respectively). Interform reliabilities were lower for Groups 2 and 3, where reliabilities were computed between a standard-time and an extended-time SAT I administration, yet correlations were still substantial, indicating rank ordering of students is not likely to vary substantially due to extended time. The extent of score gain is substantially higher for this group, indicating that there is a benefit from additional amounts of extended time. The correlation found for the verbal scale is noticeably higher than the .76 correlation reported by Centra (1986) for learning disabled students.

TABLE 9

Correlations Between Timed and Untimed SAT I Scores for Other Three Groups

	Correlations Math Score	*Correlations Verbal Score*
Group 1 (standard – standard) n = 706,537	.887*	.878*
Group 2 (extended – extended) (Learning Disabled) n = 4,479	.877*	.852*
Group 3 (standard – extended) (Learning Disabled) n =3,827	.849*	.825*
Group 4 (extended – standard) (Learning Disabled) n = 784	.848*	.820*

*p is significant at the .01 level

TABLE 10

Correlations Between Standard-Time and Extended-Time SAT I Scores for Learning Disabled Students (Group 3)

Length of Time on Untimed Administration (in minutes)	*Correlations. Math Score*	*Correlations. Verbal Score*
≤ 0	.844* n = 1,564	.812* n = 1,281
1 – 38	.866* n = 1,408	.852* n = 1,329
39 – 75	.842* n = 574	.848* n = 827
≥ 76	.801* n = 281	.783* n = 390
Total n = 3,827	.849* n = 3,827	.825* n = 3,827

*p is significant at the .01 level

Another method of examining comparability of scores resulting from different administrations is to look at the pattern of correlations as a function of additional time to complete the SAT I. To the extent that additional time affects students' scores, correlations between standard-time and extended-time testing should be reduced as students take more time to complete the test. Table 10 illustrates that interform correlations ranged between .783 and .866 and were all statistically significant (p >.01). However, the variations in correlations may be due to a more restricted range of scores, and there are few practical differences between interform correlations. The one exception may be the lower correspondence between test scores for students completing tests in more than double time (more than 76 minutes). Again, these findings support analyses conducted by Centra (1986) showing little variation among correlations between SAT scores as a function of extended time until double time is permitted.

Discussion

As predicted, the opportunity to retest with extended time allowed learning disabled students to improve their

SAT I test performance between their junior and senior years in high school. The mean score gains for learning disabled students completing a standard-time SAT I followed by an extended-time SAT I are 45 and 38 points on the verbal and math scales, respectively, as compared with mean score gains of 13 and 12 points, respectively, for nondisabled test takers completing two standard-time administrations. Adjusting for differences in score gain for these two groups provides residual gains of 32 and 26 for verbal and math. These residuals are useful in estimating the mean effects of extended time for learning disabled students across all score ranges.

These residual score gains and patterns of score gain across these groups are consistent with results reported by Centra for the SAT (1986) and Ziomek (1997) on the ACT Composite scale. Results from this study show that mean residual score gains for all students, as well as just those with initial scores in the 300–400 range—which is used in Centra's analyses (1986)—are generally consistent with his findings for the verbal scale but substantially lower for the math scale (Centra, 1986). Residual score gain was 30 on verbal in the Centra study, compared with 32 for all students, or 21 for students with initial scores between 300–400 in the present study. On the math scale, Centra reported a residual score gain of 33 points, compared to gains of 26 and 11 for all students in Group 3 and students with initial scores between 300–400, respectively.

Score gain is also slightly lower in this study than the gain reported by Ziomek (1997) on the ACT Composite scale. He reported an overall score gain of 3.2 points on the ACT Composite for special-needs students completing a standard-time ACT followed by an extended-time ACT ($n = 3{,}439$), as compared with a gain of 0.9 for nondisabled students retesting. This results in an overall residual mean score gain of 2.3; however, the ACT sample combined three groups of special-needs students, and 29 percent of these students tested with a cassette as well as receiving extended time. The average gain for students receiving extended time alone were: 4.7 for attention deficit/hyperactivity disorder, 3.2 for dyslexia, and 2.8 for learning disability. Ziomek (1997) also reported a mean score gain of 0.8 ($n = 3{,}410$) for nonspecial-needs students who tested at least twice under regular testing conditions and a mean decrease of 0.6 ($n = 437$) for special-needs students who elected to initially test under special circumstances and then retested under standard conditions. Ziomek (1997) did not attempt to examine residual gain and estimate the effects of extended time. The ACT composite ranges from 1–36, with a standard deviation of 6 (ACT, 1998).

A major problem with any analysis of the effects of accommodations for disabled examinees, such as the effects of extended time, is the difficulty in disaggregating the extent the modification compensates for the disability from the extent that it may overcompensate and introduce construct irrelevant variance (attributed to extra time to more carefully read, review, and respond to items) into the score. Extended time is provided to students with learning disabilities on the SAT I based on an assumption that increased time will result in a more precise measure of the students' achievements, implying that these students would be impeded by their disabilities if they completed a standard-time test. A second limitation with these results is that the students completing the SAT I represent a self-selected group of students who are intending to go to college, and those students requesting accommodations represent a further self-selected group of students. Students with learning disabilities requesting accommodations on the SAT I may differ in meaningful ways from other students who do not request accommodations. Further, students with disabilities completing a standard-time administration as well as an extended-time administration may also differ from students testing only with extended time.

Score gain increased in a direct relationship to the amount of extended time taken by the learning disabled students. This finding suggests that the more time a learning disabled student is allowed to use on the SAT I, the more his or her scores will increase proportionally.

This finding is logical because on the extended-time test, learning disabled students are not provided with some amount of additional time estimated to be comparable to that given to nondisabled students under standardized conditions. Instead, disabled students are able to request and receive substantially more time, which may permit them to read, respond to, and review somewhat more test items at a more relaxed pace than students testing under the standard-time limit.

Students receiving up to double time had score gains of over 100 points on SAT I math and 60 points on SAT I verbal. Approximately 10 percent of the students completing a standard-time administration first retested with more than double time (i.e., over 150 minutes) on the SAT I, and their score gains were 69.1 and 84.8 points on the math and verbal scales, respectively. However, the restricted range of scores, and other unknown differences between learning disabled students requiring different amounts of time to complete the SAT I, probably contribute to the score gain associated with additional amounts of extended time. Finally, on average, students offered extended time on the SAT I required more time for completing the verbal section than the math section.

Extended time may also slightly change the construct being measured. As substantial amounts of extended time (i.e., double time) are provided students with learning disabilities, the interform correlations are reduced. This could suggest that there is a small speededness component associated with verbal and mathematical reasoning constructs on standard-time administrations of large-scale tests such as the SAT I that are not associated with extended-time administrations. One argument is that extended-time administrations provide a more precise estimate of students' abilities than standard-time administrations, and differences in score change are more related to the effect of speed on regular administrations, rather than providing learning disabled students with any advantage due to extended time. Such an argument would suggest that nondisabled students are slightly disadvantaged because of the time limits, rather than that learning disabled students are advantaged with substantial amounts of extended time. In such an instance, the construct for standard-time administrations would differ from extended-time administrations because of a slight speededness component. However, there is substantial evidence to suggest that constructs are generally comparable, and any differences in score change are more likely to be associated with excessive amounts of additional time for some portion of learning disabled examinees.

Willingham (1988) notes that the overprediction of SAT and GRE scores for learning disabled students testing with extended time suggests that some students may be receiving more time than required to compensate for their disability and that the additional time is the major source of noncomparability. In addition, factor analytic studies also found little invariance of constructs between several groups of disabled examinees and nondisabled examinees (Rock, et al., 1988). There was some evidence that antonyms may be tapping more than a verbal reasoning construct for learning disabled students in the 1988 study; however, that item no longer appears on the SAT I. The other major finding was that verbal and mathematical reasoning appeared somewhat more independent among disabled test takers. This greater specificity of learning or ability among disabled test takers could be due to the disability or other factors such as selection bias. A more contemporary study examining the construct comparability of the SAT I is required now that changes in the content and item types were introduced with the revised test in 1994.

The relatively high significant correlations between students' timed and untimed test performance are only slightly lower than the correlations with the other three groups examined. This finding affirms that students' scores on the SAT I are highly reliable whether or not moderate amounts (up to double time) of extended time are permitted and are generally consistent with previous research in illustrating similar levels of measurement precision for disabled and nondisabled test takers under different levels of test speededness (Bennett, Rock and Kaplan, 1988; Centra, 1986).

This study also suggests that additional research is required to better explain differences in the performance of learning disabled and nondisabled students both on the test and in subsequent college courses. In addition to examining the construct comparability of the SAT I, more contemporary research examining the predictive validity of tests administered under varying amounts of extended time (and other accommodations) in predicting individual college course grades is required. Validation studies employing freshman GPA will be somewhat restricted because of potential differences in the number and level of courses learning disabled students may take during the freshman year. Differences in courses (number, content, rigor) must be controlled to detect small effects that may be likely in examining such differences among small groups. Additional research that examines the accommodations learning disabled students request and receive at college, and any effects on grades and time to graduation, would also be useful. As the number and proportion of students with learning disabilities increases, the nature and severity of their disabilities, as well as other characteristics, may differ markedly from those of disabled students receiving accommodations in the early eighties. Additional and more contemporary research is required because of the changing population of disabled test takers—brought about by public policy, technology, and improvements in measurement and assessment.

References

ACT (1998). *Technical manual for the ACT Assessment.* Iowa City, IA: ACT.

American College Testing (ACT) (January 25, 1995). Personal communications.

American Educational Research Association, American Psychological Association and National Council on Measurement in Education (1985). *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.

Asquith, N., and Feld, D. E. (1992). *Employment testing manual: 1992 cumulative supplement.* Boston: Warren, Gorman & Lamont.

Bennett, R. E., Rock, D. A., and Kaplan, B. A. (1988). Level, reliability and speededness of SAT scores for nine handicapped groups. *Special Services in the Schools, 4* (No. 3–4), 37–54.

Braun, H., Ragosta, M., and Kaplan, B. (1988). Predictive validity. In Willingham, W., Ragosta, M., Bennett, R. E., Braun, H., Rock, D. A., and Powers, D. E., *Testing handicapped people* (109–132). Boston: Allyn and Bacon.

Bridgeman, B., and Schmitt, A. (1997). Fairness issues in test development and administration. In W. Willingham and N. Cole, *Gender and fair assessment.* Mahwah, NJ: Lawrence Erlbaum.

Centra, J. A. (1986). Handicapped student performance on the Scholastic Aptitude Test. *Journal of Learning Disabilities, 19* (6), 324–327.

College Board. (1997). *Handbook for the SAT Program 1997-98: Admissions Staff.* New York: College Board.

Dana, P., and Ziomek, R. L. (1996). *Extended time testing on ACT or SAT: What difference does it make?* Paper presented at the National Association of College Admissions Counselors, Minneapolis, MN.

Donlon, T. F. (1984). *The College Board technical handbook for the scholastic aptitude test and achievement tests.* New York: College Board.

Educational Testing Service (ETS) (1996). *Plan A study 1995–96.* Princeton, NJ: ETS.

Evans, F. R. (1980). A study of the relationship among speed and power aptitude test scores, and ethnic identity. *College Board Research and Development Report, RDR 80–81, No. 2.* Princeton, NJ: ETS.

Evans, F. R., and Riley, R. R. (1972). A study of speededness as a source of test bias. *Journal of Educational Measurement, 9,* 123–31.

Evans, F. R., and Riley, R. R. (1973). A study of test speededness as a potential source of test bias in the quantitative score of the Admission Test for Graduate Study in Business. *Review in Higher Education, 1,* 173–183.

Eyde, L. D., Nester, M. A., Heaton, S. A., and Nelson, A. V. (1994). *Guide for administering written employment examinations to persons with disabilities.* Washington, DC: U.S. Office of Personnel Management, Personnel Research and Development Center.

Klein, S. P. (1981). *The effects of time limits, item sequence and question format on applicant performance on the California Bar Examination.* San Francisco: Committee of Bar Examiners of the State Bar of California and National Conference of Bar Examiners.

Lyu, C. F., and Lawrence, I. (1998). *Test-taking patterns and average score gains for the SAT.* (Unpublished statistical College Board Report No. SR-98-05). Princeton, NJ: ETS.

Mehrens, W. A. (In press). Flagging test scores: Policy, practice and research. Washington, DC: American Psychological Association.

Nathan, J. S., and Camara, W. J. (1998). Score change when retaking the SAT. *Research Note, RN 98-05.* New York: College Board.

Ragosta, M. and Wendler, C. (1992). Eligibility issues and comparable time limits for disabled and nondisabled SAT examinees. *College Board Research Report No. 92-5.* New York: College Board.

Rock, D. A., Bennett, R. E., Kaplan, B. A., and Jirele, T. (1988). Construct validity. In Willingham, W., Ragosta, M., Bennett, R. E., Braun, H., Rock, D. A., and Powers, D. E. (1988). *Testing handicapped people* (pp. 99–107). Boston: Allyn and Bacon.

Sherman, S., and Robinson, N. (Eds.). (1982). *Ability testing of handicapped people: Dilemma for government, science, and the public.* Washington, DC: National Academy Press.

Swineford, F. (1956). *Technical manual for users of test analysis.* (Statistical Report 56-42). Princeton, NJ: ETS.

Welsh, M. (1995). Personal correspondence.

Wightman, L. F. (1993). Test takers with disabilities: A summary of data from special administrations of the LSAT. *Research Report 93-03. Law School Admissions Council.* Newtown, PA: LSAT and Law School Admissions Services.

Wild, C. L., and Durso, R. (1979). *Effects of increased test-taking time on test scores by ethnic and gender subgroups.* (GRE 76-06R). Princeton, NJ: ETS.

Willingham, W., Ragosta, M., Bennett, R. E., Braun, H., Rock, D. A., and Powers, D. E. (1988). *Testing handicapped people.* Boston: Allyn and Bacon.

Ziomek, R. L. (1997). "Achievement score gains of specific-needs tested students who tested at least twice on the ACT Assessment." Unpublished paper. Iowa City, IA: ACT.

Appendix

Table A1

Percent of Learning Disabled Students With Score Changes on the SAT I Verbal Section When Testing With Extended Time in Junior and Senior Years (Group 2) (n=4,479)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				3	7	12	12	20	14	17	15		139
280–320		1	6	4	8	14	24	17	10	8	8		187
330–370	1	0	3	7	10	17	20	19	14	7	2		477
380–420		1	2	7	11	24	24	18	8	3	2		753
430–470		2	3	8	15	20	24	18	8	2			853
480–520		1	2	9	18	23	25	13	7	2			851
530–570		2	3	8	21	24	23	14	4	1			623
580–620		1	3	9	19	25	24	12	4	2	1		388
630–670		2	6	8	17	30	19	12	5		2		146
680–720		6	10	13	19	13	27	8	4				52
730–770				50	33	17							6
780–800			25		50		25						4

Table A2

Percent of Learning Disabled Students With Score Changes on the SAT I Math Section When Testing With Extended Time in Junior and Senior Years (Group 2) (n=4,479)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				3	4	7	14	18	20	15	19	311	115
280–320		2	2	5	10	13	21	19	17	8	3	340	217
330–370	1	1	4	8	12	18	24	16	10	4	2	363	561
380–420	1	1	3	7	15	23	23	15	8	3	1	411	851
430–470	1	1	2	7	14	23	23	16	8	3	2	459	851
480–520		1	2	6	16	22	21	16	10	4	2	510	751
530–570		1	3	6	16	22	20	16	9	4	3	558	493
580–620		1	4	7	15	23	20	16	9	3	2	606	368
630–670		2	3	9	19	22	21	13	7	3	1	655	191
680–720		1	3	14	20	22	16	14	6	4		707	63
730–770		6	7		25	25	31	6				740	14
780–800	14			43	29		14					782	4

Table A3

Percent of Learning Disabled Students With Score Changes on the SAT I Verbal Section When Testing With Extended Time in Junior Year and Standard Time in Senior Year (Group 4) (n=784)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				6	3	3	6	40	18	12	12	306	33
280–320		2	2	9	13	33	18	13	6	2	2	31	46
330–370	1	1	5	10	14	23	18	11	8	5	4	364	114
380–420	1	4	9	10	18	25	7	3				389	158
430–470	1	4	2	13	29	27	16	7	1			435	142
480–520	3	6	14	20	17	24	9	5	1	1		485	128
530–570	4	15	4	18	18	15	15	11				516	89
580–620	7	11	2	23	23	25	7	2				559	43
630–670	4	15	4	18	18	15	15	11				620	27
680–720		25	25	50								680	4
730–770													0
780–800													0

Table A4

Percent of Learning Disabled Students With Score Changes on the SAT I Math Section When Testing With Extended Time in Junior Year and Standard Time in Senior Year (Group 4) (n=784)

Junior-Year Scores	*≤ -140*	*-130 to -110*	*-110 to -80*	*-70 to -50*	*-40 to -20*	*-10 to 10*	*20 to 40*	*50 to 70*	*80 to 100*	*110 to 130*	*≥ 140*	*Mean Senior Score*	*N*
200–270				3	10	3	10	27	27	7		304	30
280–320			4	11	13	24	23	14	8	1	2	319	45
330–370		5		5	11	18	33	20	6	2		363	107
380–420	1	3	4	12	22	24	18	9	5	1	1	385	196
430–470		1	7	17	17	26	20	11	1			441	139
480–520	3	4	4	10	31	27	13	6	2			483	111
530–570		3	16	14	90	22	13	2	4	1		531	77
580–620		12	19	16	16	24	5	5	3			611	42
630–670		10	7	28	25	18	4	4	4			615	28
680–720			40	20	20	20						664	5
730–770					50		50					745	2
780–800					100							755	2

TABLE A5

Number and Percent of Students With Disabilities Retesting on the SAT I During Spring of Junior Year and Fall of Senior Year (April 1994–December 1995) by Accommodation

Disability	*Missing*	*Regular Print*	*Large Print*	*Braille*	*Cassette*	*Script/Reader*	*Total*
Visual	16 (8)	45 (21)	103 (49)	6 (3)	23 (11)	17 (8)	**210 (2.8)**
Physical	12 (7)	113 (68)	11 (7)	—	14 (8)	15 (9)	**165 (2.2)**
Hearing	4 (5)	57 (75)	—	—	7 (9)	8 (10)	**76 (2.0)**
Learning	202 (3)	5,331 (75)	163 (2)	—	979 (14)	476 (7)	**7,151 (94.1)**
Total	**234 (3)**	**5,546 (73)**	**277 (4)**	**6 (0)**	**1,023 (13)**	**516 (7)**	7,602

CollegeBoard

connect to college success™